DNA sequence variation and methylation in an arsenic tolerant earthworm population

Peter Kille[1#], Jane Andre[1,2#], Craig Anderson[1,3], Hui Na Ang[2,3], Michael W. Bruford[1], Jacob G. Bundy[4], Robert Donnelly[1], Mark E. Hodson[2†], Gabriela Juma[1], Elma Lahive[3], A. John Morgan[1], Stephen R. Stürzenbaum[5], David J. Spurgeon[3*]

[#] PK and JA should be considered joint first authors.

[1] Cardiff School of Biosciences, BIOSI 1, University of Cardiff, P.O. Box 915, Cardiff, CF10 3TL, UK.

[2] Soil Research Centre, School of Human and Environmental Sciences, University of Reading, Reading, RG6 6DW, Berkshire, UK.

[3] Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Wallingford, Oxfordshire OX10 8BB, UK.

[4] Biomolecular Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, Sir Alexander Fleming Building, South Kensington, London, .

[5] King's College London, Analytical and Environmental Sciences Division, 150 Stamford Street, London SE1 9NH, UK.

[*] Author for correspondence. Full contact details: Dr David Spurgeon, Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Wallingford, Oxfordshire, UK, OX10 8BB, Tel: + 44 1491 772 208; Fax: +44 1491 692 424; Email: dasp@ceh.ac.uk

[†] Current address: Environment Department, University of York, Heslington, York, YO10 5DD

Running title: DNA variation and methylation in an arsenic tolerant earthworm population

Abstract

Evidence is emerging that earthworms can evolve tolerance to trace element enriched soils. However, few studies have sought to establish whether such tolerance is determined through adaption or plasticity. Here we report results from a combined analysis of mitochondrial (cytochrome oxidase II, COII), nuclear (amplified fragment length polymorphism, AFLP) variation and DNA methylation in populations of the earthworm *Lumbricus rubellus* from sites across an abandoned arsenic and copper mine. Earthworms from the mine site population demonstrated clear arsenic tolerance in comparison to a naïve strain. COII and AFLP results suggest that *L. rubellus* from the unexposed and the adapted populations comprises two cryptic lineages (Linages A and B) each of which was present across all of the sites. AFLP analysis by lineage highlighted variations associated with soil metal/metalloid concentrations (most clearly for Lineage A) suggesting a genetic component to the observed tolerance. The methylation sensitive AFLP (Me-AFLP) identified a high genome methylation content (average 13.5%) in both lineages. For Lineage A, Me-AFLP analysis did not identify a strong association with soil arsenic levels. For Lineage B, however, a clear association of methylation patterns with soil arsenic concentrations was found. This suggests that Lineage B earthworms utilise epigenetic mechanisms to adapt to the presence of contamination. These fundamentally different genetic adjustments in the two clades indicate that the two lineages employ distinct adaptive strategies (genetic or epigenetic) in response to arsenic exposure. Mechanisms driving this variation may be founded within the colonisation histories of the lineages.

Keywords: arsenic tolerance, cryptic lineages, adaptive variation, DNA methylation, epigenetics

1. Introduction

Many invertebrate species are able to maintain viable populations in polluted soils where total and potentially bioavailable metal/metalloid concentrations greatly exceed toxicity values (e.g. LC_{50}s) for known naïve (and so sensitive) populations (laboratory strains). This suggests that under trace element exposure, some invertebrate populations develop metal tolerance through behaviour or physiological adaptive traits (Posthuma and Van Straalen, 1993; Van Straalen and Roelofs, 2007). Mechanisms underpinning this tolerance have in some cases been shown to involve heritable changes in coding or promoter regions of metal efflux pumps (Callaghan and Denny, 2002) and thiol-rich peptides involved in sequestration (glutathione-S-transferases, phytochelatins, and metallothioneins) (Janssens et al., 2007; Vatamaniuk et al., 2005). In other cases, however, the mechanisms underlying tolerance remain unknown and/or unstudied.

For earthworms, one of the most functionally important of soil taxa (Lavelle et al., 1997), indirect evidence for metal tolerance is provided by the fact that earthworms can be collected from soils containing residue levels that significantly exceed toxic effect concentrations (Spurgeon and Hopkin, 1999a, b). However, difficulties in extrapolating toxicity data between the laboratory and field due to, for example, contaminant aging and speciation (Arnold et al., 2003; Arnold et al., 2007), mean that the mere presence of earthworms in these polluted soils is not alone confirmation that adaptation has occurred (Davies et al., 2003; Donner et al., 2010). More directly in relation to tolerance, studies with successive generations of *Eisenia fetida* selected for tolerance to Zn over two generations found changes in the shape of concentration response relationships for survival that were indicative of tolerance development (Spurgeon and Hopkin, 2000). In the field, Langdon et al. (1999) noted that *L. rubellus* living in arsenic and copper polluted soil at two abandoned arsenic mines (Devon Great Consols, Carrock Fell) could survive in arsenic-spiked soil that was acutely toxic to earthworms from a clean site. This tolerance was preserved when the mine populations were reared on clean soil over two generations, suggesting a genetic basis for this phenotype (Langdon et al., 2009).

Despite indications of trace metal and metalloid tolerance in earthworms, the extent to which there is a genetic and/or physiological basis of this trait has not been fully investigated. A study of isozyme specific polymorphisms within *L. rubellus* populations known to be adapted to combined metal and flooding stress

82 failed to identify adaptive variation (Simonsen and Klok, 2010), although the results of this study should be
83 treated with some caution as enzymes known to be related to metal tolerance were not targeted. The only
84 study that has, to date, identified a potential genetic basis for tolerance to adverse soil conditions in
85 earthworms is that for *L. rubellus* living at a lead/zinc mine located at Cwmystwyth, mid Wales. For this
86 population, Andre et al. (2010) used mitochondrial (COII) and amplified fragment length polymorphism
87 (AFLP) genotyping to demonstrate that a mine spoil associated population showed little genetic overlap (in
88 AFLP profile) with individuals within populations at two less polluted sites.
89
90 While the assumption often is that individual/population survival is based on selection for increased
91 tolerance, there is evidence emerging that the plastic responses driven by chemical influences on the
92 epigenome may also be an important mechanism of adaptation (Mirouze and Paszkowski, 2011; Ren et al.,
93 2011; Seong et al., 2011). Among the many epigenetic mechanisms, DNA methylation represents a key
94 response. Given that earthworms have been recorded to possess a 13% methylated cytosine content in DNA
95 (Regev et al., 1998), the potential for mediation of adaptive tolerance through epigenetic DNA methylation
96 should be considered. Here, we report a combined toxicological and genetic study, using mitochondrial
97 (COII) and nuclear (AFLP) and DNA methylome analysis, for the earthworm *L. rubellus* sampled at sites of
98 different metal pollution status within an As-contaminated mine complex - the Devon Great Consols (DGC)
99 site in the UK. For the study, we sampled earthworms from a number of sites within DGC including one
100 (Site 2 in this study) from which tolerant populations previously studied by Langdon et al. (2009; 1999) were
101 collected. Adjacent and distant reference sites were also sampled. That arsenic, a major contaminant at the
102 site, has been reported to induce epigenetic changes including hypo-and hyper-methylation of DNA (Ren et
103 al., 2011; Zhong and Mass, 2001) makes the site particularly suited for the analysis of DNA methylation
104 responses in earthworms. Initially the collected populations were screened to confirm that the tolerance
105 previously reported for populations at study Site 2 was applicable to earthworms inhabiting this and other
106 collection sites located in the mining area. Genetic analyses were then undertaken using these populations.
107 The hypothesis tested was that *L. rubellus* populations from polluted sites that show evidence of tolerance,
108 would include individuals with mitochondrial or nuclear genotypes and/or DNA methylation patterns that
109 were distinct from those of intolerant earthworms from (adjacent) unpolluted locations.

2. Materials and Methods

2.1 Site description, sampling and soil characterisation: This study was conducted at the abandoned Devon Great Consols mine complex located in the Tamar Valley, Devon, South-West UK (UK Ordnance Survey. Map coordinates for mine centre: SX426733 – N50:32:52 W4:13:25). This mine was worked for copper and arsenic from 1844-1900 and from 1915-1930. Across the site, the spoil from various extraction processes remain. The soils established on these wastes contain highly elevated concentrations of trace elements, including arsenic and copper. Earthworms (*L. rubellus*) were sampled from six locations in the region of the Devon Great Consols mine. Four locations (Sites 1-4) were situated on the mine and waste handling area (see Fig. 1). This included a location (Site 2 i.e. close to the area where arsenic was processed using the calciner method) from which the adapted population studied by Langdon et al. (2009; 1999) was collected. Two clean reference site populations were also sampled. These were at a site adjacent to the contaminated area, but which itself was not greatly enriched in arsenic and copper (Site Control - SC) and a site some 20 km distant from DGC which was outside the geological area of arsenic rich soils present in the Tamar Valley (Off-Site Control - OSC) (UK Ordnance Survey. Map coordinates SX 418901 N50:68:89 W4:24:03).

At each site, approximately 30 fully clitellate adult *L. rubellus* were collected by digging and hand-sorting over two consecutive days in September 2010. To ensure that genome methylation patterns were not influenced by handling stress, all earthworms were washed and blotted dry on-site and then snap frozen in liquid nitrogen. Triplicate soil samples from surface to 5 cm depth were also collected from each location. These were subsequently oven dried at 80^{o}C and sieved through a 2 mm mesh to remove large roots and stones. Total concentrations of arsenic, barium, calcium, cadmium, chromium, copper, iron, magnesium, nickel, strontium and zinc were determined in a 1 g sample of the processed soil following an aqua regia digestion protocol (Arnold et al., 2008). Digests were analysed on a Perkin Elmer Optima 7300 DV inductively coupled plasma optical emission spectrometry instrument. For quality control, an in house reference traceable to BCR-143R (Commission of the European Communities, Community Bureau of Reference) was included with each batch of digestions. Measured concentrations were always greater than 75% of reference values and were above 95% for As. Organic matter content of each soil sample was

137 measured by proxy using loss on ignition following combustion at 500°C (Rowell, 1994) and pH was
138 quantified by electrode from a 1:5 volume soil:water mix (International Standards Organisation, 2005).
139
140 *2.2. Toxicity tests to identify putative arsenic tolerance*
141 To identify potential tolerance, a 14 day exposure to a single pre-determined arsenic concentration was
142 undertaken to compare survival patterns of earthworms from sites located within and adjacent to the
143 DGC complex to those for a known naïve population. The soil concentration used for this assay was
144 derived from a preliminary study conducted to assess survival of the naïve population at 150 and
145 300 mg/kg arsenic. The earthworms used were taken from a culture established from a field
146 collected population (Lasebo BV, Nijkerkerveen, The Netherlands). At each tested concentration,
147 15 replicate containers, each including 200 g dry weight of a clay loam soil (Broughton Loams,
148 Kettering, UK) (see Spurgeon et al., 2003), were spiked with sodium arsenate solution (Santa Cruz
149 Biotechnology Inc., Santa Cruz, California, US) to give the required metalloid concentration and a
150 soil moisture content of approximately 50% of field capacity. After a one week stabilisation period,
151 one adult *L. rubellus* was added to each replicate and kept at 13 ±1 °C under constant light for seven
152 days. Earthworms were observed daily and mortality recorded. Based on these findings, a screening
153 concentration of 300 mg/kg arsenic was selected for the definitive tolerance assay, since this
154 concentration resulted in progressive mortality of the naïve earthworms over the exposure period.
155 Thus the definitive assay was conducted using the 300 mg/kg concentration with 15 earthworms
156 from each of the DGC sites (Sites 1-4) and the SC reference population. The exposure was extended
157 to 14 days to allow the potential to identify survival patterns in more tolerant populations.
158
159 *2.3 Mitochondrial cytochrome oxidase II (mtCOII)) sequencing:*
160 DNA was purified from ~10 mg of tissue from the anterior of each individual using the DNAzol reagent
161 (Life Technologies, Paisley, UK). PCR amplification of the cytochrome oxidase II (COII) mitochondrial
162 gene made use of forward (TAGCTCACTTAGATGCCA) and reverse (GTATGCGGATTTCTAATTGT)

163 primers and was conducted following Andre et al. (2010). PCR products were assessed electrophoretically
164 prior to purification and sequencing using ABI PRISM® BigDye v3.1 Terminator technology (Applied
165 Biosystems, USA). Obtained sequences were aligned by ClustalW prior to tree construction using the
166 Maximum Likelihood (ML) and Bayesian methods in Mega v5.01 and MRBAYES v3.2, respectively. ML
167 estimation incorporated the Tajima-Nei model, supported by bootstrap analyses over 1000 iterations.
168 Bayesian analyses were conducted using a General Time Reversible model with a proportion of invariable
169 sites and a gamma-shaped distribution over 2 independent runs. Four Markov Chains were run over 2 million
170 iterations and sampled every 1000 generations, with the first 500 trees discarded as burn-in. Both
171 phylogenetic estimates incorporated outlier sequences from *Lumbricus castaneus* and *Lumbricus terrestris* as
172 well as sequences that represent previously recognised *L. rubellus* clades (Andre et al., 2010).
173
174 *2.4 AFLP and methylation sensitive AFLP profiling*: A combined AFLP and Me-AFLP protocol was
175 optimised in a pilot methylation study and was based on parallel use of methylation- and non-methylation-
176 sensitive restriction enzymes (*Hap*I and *Msp*I) to treat DNA samples prior to primer ligation and
177 amplification (Xiong et al., 1999). Both *Hap*II and *Msp*I recognize a CCGG sequence; however, while *Msp*I
178 is able to cut methylated recognition sites (as well as unmethylated ones), *Hap*II is unable to cut at such
179 locations when they are methylated (i.e. only unmethylated recognition sites are cut). The extent of
180 methylation of restriction sites can therefore be ascertained by recording bands amplified by *Msp*I but not
181 *Hap*II. Such bands can be used to compare individual methylation patterns. AFLP analysis was conducted
182 for individuals from the six collection locations using pre-selective primers and analysis on an Applied
183 Biosystems 3130 x 1 fragment analyser (Andre et al., 2010). Cumulative AFLP fragment profiles were
184 transformed to a binary form and principal coordinates (PCO) analysis used to visualise the genetic
185 relationship between individuals using GenAlEx 6.4.1.
186

187 ## 3. Results

188 Soil analyses highlighted the extent and severity of the arsenic (and copper) contamination at DGC. Arsenic
189 and copper levels were greatly elevated in soils from all sites on the mining area (Sites 1-4), with cobalt also
190 higher than SC and OSC soils by at least a factor of two at the sites (Table 1). The most polluted arsenic soil
191 (Site 4) contained almost 20,000 mg/kg of arsenic, over 900 mg/kg copper and also elevated cobalt,
192 cadmium and lead levels. The remaining three mine spoil contaminated sites each contained over 4000
193 mg/kg As and over 500 mg/kg copper. As expected, the lowest concentrations of arsenic and copper and
194 other trace metals were found at the SC and OSC reference sites. Levels at SC were in the 300 mg/kg As
195 range, a concentration still elevated above background arsenic levels for British soils (Emmett et al., 2010).
196 OSC soils contained arsenic levels consistent with these background concentrations.
197
198 Measured site soil LOI and pH values are presented in Table 1. Whilst the pH of all four sites located on the
199 mine area and the OSC was similar acidic (pH 4.1 - 4.8), the SC site had a pH of 5.6. Overall, there was
200 minimal pH variation between sites, and no correlation with soil arsenic or copper concentration (Pearson
201 correlation, $p > 0.05$). The absence of a correlation indicates that soil pH influences are unlikely to confound
202 attempts to link genetic variation to soil contaminant levels. For LOI, the lowest values in the mining site
203 soils (4.2 - 17.2%) were found at Sites 1 and 4, while the remaining two soils had higher LOI values (29.7 –
204 49.6%). This may reflect the vegetation of the sites: open in the case of Sites 1 and 4, wooded at Sites 2 and
205 3. The two control site soils had %LOI levels intermediate within the range of the two pairs of mine
206 sampling locations.
207
208 Exposure of the naïve population to 300 mg/kg of arsenic in soil resulted in a progressive mortality,
209 culminating in only 7% survival after 14 days of exposure (Fig. 2). In the SC population progressive
210 mortality over time was also seen. This, however, proceeded at a slower rate than for the naïve earthworms,
211 culminating in 46% survival after 14 days. These variations in mortality rates resulted in different LT_{50}
212 estimates from Weibull models fits (SigmaPlot 12.0) for the naïve and SC populations; these being 5.3 (95%
213 Confidence Intervals 4.9-5.6) and 12.4 (95% Confidence Intervals 11.6-13.3) days respectively. In the four
214 DGC mine site populations there was observable mortality in the Site 1 population, although 73% survival

after 14 days was higher than for either the naïve or SC earthworms. Populations from the remaining three DGC mine sites show low mortality, with 100% survival for Site 2 and 4 earthworms and 85% survival for Site 3 earthworms.

The mtDNA COII analysis indicated the presence of two distinct lineages (A and B) within the sampled *L. rubellus* (Fig. 3a). The two cryptic lineages show a 18% and 14% genetic divergence from *L. castaneus* and *L. terrestris* respectively. Average difference between lineages was 10.3%. Internal within the lineages, the Lineage A earthworms have a maximum 1.4% genetic difference, while for Lineage B earthworms this was 0.06%. This high level of divergence between the two major lineage branches identifies *L. rubellus* as a complex of cryptic lineages as found previously (Andre et al., 2010). A comparison of the frequency of lineage occurrence at each sampled site identified differences in population COII haplogroup composition. Populations at two sites, Site 4 and Site OSC, included 90% or more of individuals from Lineage B; while in contrast the Site 3 populations included 76% of Lineage A individuals. The remaining three sites each had an approximately equal proportion of each lineage, with Lineage A slightly dominant at Site SC (57%) and Site 1 (54%) and Lineage B at Site 2 (64%). That both lineages were found at all sites, often in similar proportions, and also that the two sites showing greatest lineage selection (dominance of Lineage B at both Site 4 and Site OSC) included both the most and least arsenic contaminated soils, is indicative of an absence of a mitochondrial lineage association with soil contamination status.

Standard AFLP analysis conducted using *MspI* (which cuts at all recognition sites independent of methylation status) reemphasised the presence of two *L. rubellus* lineages as indicated by the mitochondrial markers. All mine site earthworms fell clearly into one of the two major lineages, but apparent inter-lineage individuals were observed among SC and OSC earthworms. These hybrids show AFLP genotypes intermediate between the two lineages on PC1 and divergent on PC2 (Fig. 3b). The presence of hybrids is in agreement with previous observation of AFLP profiles in *L. rubellus* (Andre et al., 2010). The dominance of the lineage effects within a PCO analysis of the AFLP data meant it was not possible to visualise site effects within the complete data-set. Consequently independent lineage-based analyses were conducted (n.b. putative hybrid individuals were excluded from these analyses).

243

244 Within Lineage A, PCO highlighted a site dependent effect on AFLP marker patterns. Within the PC1 and
245 PC2 scores plot, SC earthworms were clearly separated from earthworms collected from Sites 2 and 3, with
246 the Site 1 individuals intermediate and closer to the SC earthworms (Fig. 4a). Scores for PC2 (and also PC3),
247 but not PC1 within the PCO were significantly correlated with site soil arsenic concentration (Pearson
248 correlation $p < 0.01$). This significant association suggests that soil arsenic concentrations, as well as
249 possibly the concentration of other co-correlated metals such as Cu, are an important driver of genome
250 structure in Lineage A *L. rubellus* across the DGC site and surrounding area.

251

252 For Lineage B PCO analysis did not identify a separation of populations within a PC1 and PC2 scores plot,
253 although a partial separation of Site 1 and 4 was evident (data not shown). Both of these populations,
254 however, overlap with profiles from the SC and OSC earthworms within this plot. Correlation of PC1 and
255 PC2 scores with soil arsenic concentrations were not significant. Only for the PC3 score was a significant
256 correlation found (Pearson correlation $p < 0.02$) indicating a weak separation underpinned by the distribution
257 particularly of the Site 2 and Site 4 individuals on this component (Fig. 4b). These results identify that while
258 soil metals such as arsenic and correlated elements are a driver for genome structure in Lineage B, these
259 factors are less important than for Lineage A with effects only observed for the lower contribution PCs.

260

261 To assess the patterns of genome methylation in individual earthworms, a second AFLP analysis was
262 conducted using the *Msp*I methylation sensitive restriction enzyme. Me-AFLP indicated that the genome of
263 *L. rubellus* had an approximate 13.5% methylated cytosine (m5C) residue content. Across the mine sampling
264 locations, the average extent of genome methylation ranged from 10.6% in earthworms at Site 1 to 22.1% for
265 Site 4. Even though the highest average genome methylation content was at the most arsenic polluted site,
266 the fact that earthworms from the two reference sites had intermediate average methylation levels (SC
267 19.4%, OSC 13.2%) meant there was no clear correlation (Pearson correlation $p > 0.05$) between methylation
268 level and soil arsenic concentration. This suggests that in the mine soils, arsenic does not have a strong
269 global hyper- or hypo-methylation effect for the resident earthworms.

270

271 Within the Me-AFLP analysis the presence of two distinct *L. rubellus* lineages was reconfirmed.
272 Consequently, lineage-specific Me-AFLP profiles were analysed, with the hybrid individuals excluded. For
273 the Lineage A PCO analysis, a segregation of individuals collected from Site 2 and Site 4 was identified
274 within the PC1 and PC2 score plot (Fig. 4c). The remaining sites showed substantial overlap between
275 individual profiles. Correlations of PC1, PC2 and PC3 scores with measured soil arsenic concentration were
276 non-significant in all cases (Pearson correlation $p > 0.05$). This suggests that soil arsenic was not the
277 principal driver of methylation pattern difference between individuals. For Lineage B, there was a partial
278 separation of profiles of earthworms from the SC and OSC locations from individuals collected from each
279 sampled mine site population (Fig. 4d). Correlation of PC1, PC2 and PC3 scores with measured site soil
280 arsenic concentration indicated a significant correlation for the first principle component (Pearson correlation
281 $p < 0.02$). This indicates that soil arsenic (and co-correlated trace metals) represents a potentially significant
282 driver of earthworm genome methylation status for lineage B earthworms.
283
284

285 # 4. Discussion

286 Soil contamination by mineral extraction, fossil fuel consumption, waste disposal and pesticide use is a
287 common problem (Hall et al., 2006). Among trace elements, arsenic represents one of the greatest hazards
288 because of its widespread distribution and toxicity to humans and wildlife (Chen et al., 2009; Thomas et al.,
289 2001). The toxicity of arsenic has been established for earthworms. Meharg et al. (1998) determined an LC_{50}
290 of approximately 100 mg/kg As for *Lumbricus terrestris* after 8 days and Fischer and Koszorus (1992) found
291 that a 25 mg/kg potassium arsenate exposure reduced growth and cocoon production in *Eisenia fetida*. For *L.*
292 *rubellus*, Langdon et al. (2001) found an LC_{50} of 96 mg/kg As for a clean site population, although
293 populations from mine sites (including DGC) gave higher values (up to 1,510 mg/kg) suggesting tolerance.
294 Building on this work, Langdon et al. (2009) revealed that the adaptation in the mine site earthworms was
295 maintained when earthworms were bred for two generations on clean soil. Cross-tolerance to copper was also
296 found (Langdon et al., 2001).
297
298 In the test to assess the presence of potential tolerance in *L. rubellus* collected from the DGC line complex
299 area sites, there was a clear indication that the populations inhabiting the DGC site locations substantially
300 enriched in arsenic display a tolerance phenotype. Earthworms sampled from the populations at Site 1-4 all
301 showed low mortality on exposure to a soil arsenic concentration that induced acute toxicity in earthworms
302 from a naïve population and also in the SC reference population. Interestingly the different rates of mortality
303 in naïve and SC earthworms, as highlighted by differences in LT50s for these populations suggest that SC
304 earthworms possess a partial arsenic tolerant phenotype. This may be related to the presence of arsenic
305 concentrations that greatly exceed accepted background concentrations in SC soil (Emmett et al., 2010).
306
307 Tolerance to chemical exposure can classically take two forms. Most simply, it can be the result of
308 phenotypic plasticity. In this case, exposure to a substance upregulates biochemical pathways (e.g. metal
309 binding proteins, mono-oxygenases and multi-drug resistance transporters), which work to detoxify or
310 eliminate the substance. If the exposure is removed, upregulation of detoxification systems can persist,
311 predisposing individuals to deal with a future chemical challenge. This plasticity has been widely reported in
312 human subjects subjected to long-term drug exposure (Stewart and Badiani, 1993) and also in species

313 exposed to toxicants in the field (Rajamohan and Sinclair, 2009; Romach et al., 2000). Maintenance of
314 elevated protein levels and the widely reported effects of stressor exposure on the epigenome (Martinez-
315 Zamudio and Ha, 2011), including arsenic (Ren et al., 2011), can provide a mechanism through which such
316 tolerance may be temporally conserved.
317
318 A second mechanism of tolerance development exploits adaptive variation within populations. There is good
319 evidence that this kind of adaptive selection can occur in response to long term chemical exposure. One
320 example is driven by the selection of alleles coding for amino acids associated with active sites of
321 detoxification enzymes. Pesticide resistance is frequently underpinned by this mechanism, with polymorphic
322 cytochrome P450 genes often the selection target (Karunker et al., 2008; Miyo and Oguma, 2010). For
323 metals, selection for metallothionein promoter alleles and other trans-acting genetic factors has been found to
324 underpin cadmium tolerance in the collembolan *Orchesella cincta* (Janssens et al., 2007; Roelofs et al., 2006;
325 van Straalen et al., 2011).
326
327 Characterisation of metallothionein promoter alleles of earthworms collected from metalliferous and
328 unpolluted soils has so far failed to detect adaptive variation (Stürzenbaum et al., 2004). With evidence for
329 targeted selection absent, a logical next step is to move to genome wide analysis (Baird et al., 2008;
330 Hohenlohe et al., 2010). Using a combined approach applying mitochondrial genotyping and conventional
331 and methylation-sensitive AFLPs, an analysis of both genotypic and epigenetic associations of the confirmed
332 adapted and putative reference populations of *L. rubellus* with different metal/metalloid exposure histories
333 was conducted. The aim was to assess the basis of the arsenic tolerance observed in the toxicity test. The
334 mitochondrial genotyping and AFLP profiling (using both methylation sensitive and insensitive enzymes) all
335 indicated that *L. rubellus* comprises two distinct lineages that differ by over 10% in their mitochondrial COII
336 sequence. This reflects the presence of two cryptic lineages within the morphospecies (Andre et al., 2010).
337 Hybrid individuals were found although only at the two reference sites. This prevalence in uncontaminated
338 soils does not support a role of pollution in the breakdown of species boundaries as found by Vonlanthen et
339 al. (2012).
340

341 As in a previous study with *L. rubellus* from polluted landscapes (Andre et al., 2010), there was no evidence
342 of a lineage or intra-lineage haplotype association with either polluted or unpolluted sites. This supports the
343 decision to move to a more detailed analysis of population structure. The AFLP analysis conducted for
344 Lineage A indicated a clear separation of earthworms between sites, with the most important principal
345 components associated with soil pollution status. For Lineage B, an influence of soil arsenic on AFLP profile
346 was also found, albeit in this case within one of the more minor principle components (PC3). Such
347 associations that link genetic distance to pollution status have been observed in previous field studies of
348 aquatic invertebrates (Martins et al., 2009) and for both genetic units of the phylogeographically divergent
349 metallophyte *Arabidopsis halleri* (Pauwels et al., 2012). Such relationships point to a genetic component that
350 may underpin the previous observations of arsenic tolerance in *L. rubellus* collected at Site 3 by Langdon et
351 al. (2009; 1999), especially given the high frequency of Lineage A individuals at this site.
352
353 Although sequence driven differentiation between populations can clearly be important, there is emerging
354 evidence that epigenetic effects can also play a role in adaptation to local environmental conditions. Known
355 epigenetic mechanisms include DNA methylation, histone modifications, and small interfering (siRNA), and
356 micro RNAs (miRNA). Of these, DNA methylation has so far been most widely studied in animals (Suzuki
357 and Bird, 2008). Studies have identified that metals and metalloids can perturb DNA methylation including
358 hypomethylation by Cd (Takiguchi et al., 2003) and targeted gene silencing via hypermethylation by Ni (Lee
359 et al., 1995). For arsenic, the potential competition with DNA for methyl groups for respectively methyl
360 metabolites and DNA modification can create an interplay between hypomethylation (Arita and Costa, 2009;
361 Zhao et al., 1997) and hypermethylation (Jensen et al., 2008) in arsenic toxicology (Ren et al., 2011).
362
363 To date relatively little is known about the role of DNA methylation as a component of adaptive variation in
364 invertebrate organisms. Studies on a range of invertebrate species have highlighted extensive variation in the
365 5-methyl cytosine content of the genome (Regev et al., 1998). Thus, while some species, including the
366 nematode *Caenorhabditis elegans* and fruitfly *Drosophila melanogaster*, have low to negligible 5-methyl
367 cytosine levels (Bird, 2002; Regev et al., 1998), some taxa possess methylation levels in the 10-15% range.
368 Me-AFLP indicated an approximate 13.5% methylated cytosine (m5C) residue content in the *L. rubellus*

369 genome. This represents a high level of DNA methylation for an invertebrate species, but is consistent with
370 previous results for the earthworm *Aporrectodea caliginosa* (Regev et al., 1998). This suggests that DNA
371 methylation may have an important role in annelids, although to date relatively little is known about how
372 such methylation is controlled. For example, a study on the marine annelid species *Chaetopterus*
373 *variopedatus* was able to identify a protein that had a high homology to known invertebrate
374 methyltransferases, but could not confirm a role of this protein in 'de-novo' methylation of double stranded
375 DNA (del Gaudio et al., 1999).
376
377 On exposure to arsenic (and co-contaminant metals), an analysis of methylation patterns using the MeAFLP
378 approach showed a site-specific influence. For Lineage A earthworms, separation between sites for the
379 MeAFLP profiles was seen. This separation could not, however, be significantly associated with soil arsenic
380 concentration as was the case for the standard AFLP analysis for this Lineage. This may indicate that other
381 soil, biotic and local scale climatic factors may instead be modifying the epigenome. For Lineage B
382 earthworms, pattern of DNA methylation could be significantly related to soil arsenic levels, suggesting a
383 potential role of trace element exposure, although it is also feasible that environmental factors (e.g. soil
384 texture, soil moisture, food availability), that are co-correlated to soil pollutant levels, could also be
385 important. Evidence from detailed analysis of stress associated genes, such as metallothionein in the snail
386 *Helix pomatia*, has already identified the presence of genomic regions that confer a high potential for
387 epigenetic regulation indicating a potential role for epigenetic mechanisms in metal responses (Egg et al.,
388 2009). Further, in *D. melanogaster* stress exposure has been shown to result in epigenetic heterochromatic
389 disruption that is transmissible in a non-Mendelian fashion (Seong et al., 2011). The association of DNA
390 methylation patterns with arsenic exposure observed here suggests a potential role of epigenetic mechanisms
391 in stress adaptation in earthworms that concur with the evidence available for other taxa.
392
393 To extend the understanding of the role of genetic and epigenetic modification, a fruitful avenue for
394 extending this novel study from a strong associative appreciation to a mechanistic understanding of arsenic-
395 mediated molecular-genetic adaptations would entail assaying the transcription levels of specific genes
396 known to be involved in metal/metalloid trafficking and metabolism. Moreover, establishing whether

epigenetic marks are preferentially targeted to such genes and their regulatory regions in earthworms exposed to elevated levels of methylation-modifying arsenic in their native field soils is also a matter of priority. Such work clearly has the potential to link genotype to phenotype in adapted populations, so providing insight into the functional basis of adaptive traits in a key soil dwelling taxon.

The variation in lineage-specific responses observed across the genome and epigenome raises the intriguing prospect that the two *L. rubellus* cryptic lineages may employ different strategies to response to long-term arsenic exposure. The strong AFLP based separation of Lineage A earthworms in relation to soil arsenic concentrations across major PCs suggests that in this Lineage substantial genome modification has occurred as a result of long-term exposure. In contrast, the evidence for sequence modification is somewhat less compelling in Lineage B and therefore changes in genome methylation status seem to play a role in facilitating plasticity in response to soil arsenic concentration as indicated by the Me-AFLPs. Previous studies have identified differences in sensitivity between closely related lineages or species to chemical exposure. An example is the role of biotransformation capacity for determining the sensitivity of *Capitella capitata* "species" to PAH exposure (Bach et al., 2005). However, to date we are not aware of any studies that have identified such divergent genome responses to chemical exposure within two genetic lineages of a known morphospecies. The detailed basis for the evolution of distinct genetic and/or epigenetic mechanisms that drive arsenic adaptation in the two *L. rubellus* lineages, thus, emerge as potential models that could be further exploited to understand species plasticity in response to long-term chemical stress.

The genetic structure evident within the putative *L. rubellus* lineages is consistent with expectations in relation to survival within glacial refugia and subsequent recolonisation, as has been demonstrated for other species (Hewitt, 1999; Provan and Bennett, 2008). Patterns of recolonisation during the Holocene, including recent human-mediated dispersal, may have resulted in different lineages reaching the DGC area over different timeframes. Andre et al. (2010) inferred that the two *L. rubellus* lineages have very different evolutionary histories with Lineage A representing a stationary population that has experienced multiple introductions and bottleneck episodes with expansion estimated to have occurred about 250,000 years BP, while Lineage B comprises an unimodal mismatch distribution with an estimated post-glacial population

425 expansion time of approximately 17,000 years BP. It is perhaps this differential in the timescale for
426 adaptation to local arsenic contamination that has determined the lineage specific balance between adaptive
427 variation and plasticity for the two lineages at sites across the DGC mine.
428
429

Acknowledgements

This work was partly supported by the Natural Environmental Research Council grant (NE/H009973/1). We thank Helen Hooper and Huw Ricketts and others for help with sample collection and Dr Claus Svendsen for discussion of the results. We also recognise the contribution of Mr and Mrs Morgan (Snr) for the production of AJM in 1948 and at a time of the year appropriate for earthworm sampling.

436 **References**

437

438 Andre, J., King, R.A., Stürzenbaum, S.R., Kille, P., Hodson, M.E., Morgan, A.J., 2010. Molecular
439 genetic differentiation in earthworms inhabiting a heterogeneous Pb-polluted landscape.
440 Environmental Pollution 158, 883-890.

441

442 Arita, A., Costa, M., 2009. Epigenetics in metal carcinogenesis: nickel, arsenic, chromium and
443 cadmium. Metallomics 1, 222-228.

444

445 Arnold, B.E., Hodson, M.E., Charnock, J., Peijnenburg, W., 2008. Comparison of subcellular
446 partitioning, distribution, and internal speciation of Cu between Cu-tolerant and naive populations
447 of *Dendrodrilus rubidus* Savigny. Environmental Science & Technology 42, 3900-3905.

448

449 Arnold, R.E., Hodson, M.E., Black, S., Davies, N.A., 2003. The influence of mineral solubility and
450 soil solution concentration on the toxicity of copper to Eisenia fetida Savigny. Pedobiologia 47,
451 622-632.

452

453 Arnold, R.E., Hodson, M.E., Comber, S., 2007. Does speciation impact on Cu uptake by, and
454 toxicity to, the earthworm Eisenia fetida? European Journal of Soil Biology 43, S230-S232.

455

456 Bach, L., Palmqvist, A., Rasmussen, L.J., Forbes, V.E., 2005. Differences in PAH tolerance
457 between *Capitella* species: Underlying biochemical mechanisms. Aquatic Toxicology 74, 307-319.

458

459 Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U.,
460 Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced
461 RAD markers. Plos One 3, e3376.

462

463 Bird, A., 2002. DNA methylation patterns and epigenetic memory. Genes & Development 16, 6-21.
464 Callaghan, A., Denny, N., 2002. Evidence for an interaction between p-glycoprotein and cadmium
465 toxicity in cadmium-resistant and -susceptible strains of *Drosophila melanogaster*. Ecotoxicology
466 and Environmental Safety 52, 211-213.

467

468 Chen, B.W., Hua, N., Lu, M.L., Le, X.C., 2009. Metabolism, toxicity, and biomonitoring of arsenic
469 species. Progress in Chemistry 21, 474-482.

Davies, N.A., Hodson, M.E., Black, S., 2003. The influence of time on lead toxicity and bioaccumulation determined by the OECD earthworm toxicity test. Environmental Pollution 121, 55-61.

del Gaudio, R., Di Giaimo, R., Potenza, N., Branno, M., Aniello, F., Geraci, G., 1999. Characterization of a new variant DNA (cytosine-5)-methyltransferase unable to methylate double stranded DNA isolated from the marine annelid worm *Chaetopterus variopedatus*. FEBS Letters 460, 380-384.

Donner, E., Broos, K., Heemsbergen, D., Warne, M.S.J., McLaughlin, M.J., Hodson, M.E., Nortcliff, S., 2010. Biological and chemical assessments of zinc ageing in field soils. Environmental Pollution 158, 339-345.

Egg, M., Hockner, M., Brandstatter, A., Schuler, D., Dallinger, R., 2009. Structural and bioinformatic analysis of the Roman snail Cd- Metallothionein gene uncovers molecular adaptation towards plasticity in coping with multifarious environmental stress. Molecular Ecology 18, 2426-2443.

Emmett, B.A., Reynolds, B., Chamberlain, P.M., Rowe, E., Spurgeon, D., Brittain, S.A., Frogbrook, Z., Hughes, S., Lawlor, A.J., Poskitt, J., Potter, E., Robinson, D.A., Scott, A., Wood, C., Woods, C., 2010. Countryside Survey: Soils Report from 2007. NERC Centre for Ecology and Hydrology, Wallingford, UK.

Fischer, E., Koszorus, L., 1992. Sublethal effects, accumulation capacities and elimination rates of As, Hg and Se in the manure worm, *Eisenia-foetida* (Oligochaeta, Lumbricidae). Pedobiologia 36, 172-178.

Hall, J.R., Ashmore, M., Fawehinmi, J., Jordan, C., Lofts, S., Shotbolt, L., Spurgeon, D.J., Svendsen, C., Tipping, E., 2006. Developing a critical load approach for national risk assessments of atmospheric metal deposition. Environmental Toxicology and Chemistry 25, 883-890.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68, 87-112.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., Cresko, W.A., 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet 6, e1000862.

International Standards Organisation, 2005. Determination of pH., Geneva, Switzerland.

Janssens, T.K.S., Marien, J., Cenijn, P., Legler, J., van Straalen, N.M., Roelofs, D., 2007. Recombinational micro-evolution of functionally different metallothionein promoter alleles from *Orchesella cincta*. BMC Evolutionary Biology 7.

Jensen, T.J., Novak, P., Eblin, K.E., Gandolfi, A.J., Futscher, B.W., 2008. Epigenetic remodeling during arsenical-induced malignant transformation. Carcinogenesis 29, 1500-1508.

Karunker, I., Benting, J., Lueke, B., Ponge, T., Nauen, R., Roditakis, E., Vontas, J., Gorman, K., Denholm, I., Morin, S., 2008. Over-expression of cytochrome P450 CYP6CM1 is associated with high resistance to imidacloprid in the B and Q biotypes of *Bemisia tabaci* (Hemiptera : Aleyrodidae). Insect Biochemistry and Molecular Biology 38, 634-644.

Langdon, C.J., Morgan, A.J., Charnock, J.M., Semple, K.T., Lowe, C.N., 2009. As-resistance in laboratory-reared F1, F2 and F3 generation offspring of the earthworm *Lumbricus rubellus* inhabiting an As-contaminated mine soil. Environmental Pollution 157, 3114-3119.

Langdon, C.J., Piearce, T.G., Black, S., Semple, K.T., 1999. Resistance to arsenic-toxicity in a population of the earthworm *Lumbricus rubellus*. Soil Biology and Biochemistry 31, 1963-1967.

Langdon, C.J., Piearce, T.G., Meharg, A.A., Semple, K.T., 2001. Survival and behaviour of the earthworms *Lumbricus rubellus* and *Dendrodrilus rubidus* from arsenate-contaminated and non-contaminated sites. Soil Biology & Biochemistry 33, 1239-1244.

Lavelle, P., Bignell, D., Lepage, M., Wolters, V., Roger, P., Ineson, P., Heal, O.W., Dhillon, S., 1997. Soil function in a changing world: the role of invertebrate ecosystem engineers. European Journal of Soil Biology 33, 159-193.

538 Lee, Y.W., Klein, C.B., Kargacin, B., Salnikow, K., Kitahara, J., Dowjat, K., Zhitkovich, A.,
539 Christie, N.T., Costa, M., 1995. Carcinogenic nickel silences gene-expression by chromatin
540 condensation and DNA methylation - a new model for epigenetic carcinogens. Molecular and
541 Cellular Biology 15, 2547-2557.
542
543 Martinez-Zamudio, R., Ha, H.C., 2011. Environmental epigenetics in metal exposure. Epigenetics
544 6, 820-827.
545
546 Martins, N., Bollinger, C., Harper, R.M., Ribeiro, R., 2009. Effects of acid mine drainage on the
547 genetic diversity and structure of a natural population of *Daphnia longispina*. Aquatic Toxicology
548 92, 104-112.
549
550 Meharg, A.A., Shore, R.F., Broadgate, K., 1998. Edaphic factors affecting the toxicity and
551 accumulation of arsenate in the earthworm *Lumbricus terrestris*. Environmental Toxicology and
552 Chemistry 17, 1124-1131.
553
554 Mirouze, M., Paszkowski, J., 2011. Epigenetic contribution to stress adaptation in plants. Current
555 Opinion in Plant Biology 14, 267-274.
556
557 Miyo, T., Oguma, Y., 2010. Contributions of three-site mutations in acetylcholinesterase and
558 cytochrome P450 to genetic variation in susceptibility to organophosphate insecticides within a
559 natural population of *Drosophila melanogaster*. Population Ecology 52, 159-169.
560
561 Pauwels, M., Vekemans, X., Gode, C., Frerot, H., Castric, V., Saumitou-Laprade, P., 2012. Nuclear
562 and chloroplast DNA phylogeography reveals vicariance among European populations of the model
563 species for the study of metal tolerance, *Arabidopsis halleri* (Brassicaceae). New Phytologist 193,
564 916-928.
565
566 Posthuma, L., Van Straalen, N.M., 1993. Heavy-metal adaptation in terrestrial invertebrates: a
567 review of occurrence, genetics, physiology and ecological consequences. Comparative
568 Biochemistry and Physiology 106c, 11-38.
569
570 Provan, J., Bennett, K.D., 2008. Phylogeographic insights into cryptic glacial refugia. Trends in
571 Ecology & Evolution 23, 564-571.

572
573 Rajamohan, A., Sinclair, B.J., 2009. Hardening trumps acclimation in improving cold tolerance of
574 *Drosophila melanogaster* larvae. Physiological Entomology 34, 217-223.
575
576 Regev, A., Lamb, M.J., Jablonka, E., 1998. The role of DNA methylation in invertebrates:
577 Developmental regulation or genome defense? Molecular Biology and Evolution 15, 880-891.
578
579 Ren, X.F., McHale, C.M., Skibola, C.F., Smith, A.H., Smith, M.T., Zhang, L.P., 2011. An emerging
580 role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. Environmental Health
581 Perspectives 119, 11-19.
582
583 Roelofs, D., Overhein, L., de Boer, M.E., Janssens, T.K.S., van Straalen, N.M., 2006. Additive
584 genetic variation of transcriptional regulation: metallothionein expression in the soil insect
585 *Orchesella cincta*. Heredity 96, 85-92.
586
587 Romach, E.H., Zhao, C.Q., Del Razo, L.M., Cebrian, M.E., Waalkes, M.P., 2000. Studies on the
588 mechanisms of arsenic-induced self tolerance developed in liver epithelial cells through continuous
589 low-level arsenite exposure. Toxicological Sciences 54, 500-508.
590
591 Rowell, D.L., 1994. Soil Science: Methods and Applications. Longman Scientific and Technical,
592 Harlow, UK.
593
594 Seong, K.H., Li, D., Shimizu, H., Nakamura, R., Ishii, S., 2011. Inheritance of stress-induced, ATF-
595 2-dependent epigenetic change. Cell 145, 1049-1061.
596
597 Simonsen, V., Klok, C., 2010. Genetic and ecological impacts of heavy metal and flooding stress on
598 the earthworm *Lumbricus rubellus* in floodplains of the Rhine river. Soil Biology & Biochemistry
599 42, 270-275.
600
601 Spurgeon, D.J., Hopkin, S.P., 1999a. Seasonal variation in the abundance, biomass and biodiversity
602 of earthworms in soils contaminated with metal emissions from a primary smelting works. Journal
603 of Applied Ecology 36, 173-183.
604

605 Spurgeon, D.J., Hopkin, S.P., 1999b. Tolerance to zinc in populations of the earthworm *Lumbricus*
606 *rubellus* from uncontaminated and metal-contaminated ecosystems. Archives of Environmental
607 Contamination and Toxicology 37, 332-337.
608
609 Spurgeon, D.J., Hopkin, S.P., 2000. The development of genetically inherited resistance to zinc in
610 laboratory-selected generations of the earthworm *Eisenia fetida*. Environmental Pollution 109.
611
612 Spurgeon, D.J., Svendsen, C., Weeks, J.M., Hankard, P.K., Stubberud, H.E., Kammenga, J.E.,
613 2003. Quantifying copper and cadmium impacts on intrinsic rate of population increase in the
614 terrestrial oligochaete *Lumbricus rubellus*. Environmental Toxicology and Chemistry 22, 1465-
615 1472.
616
617 Stewart, J., Badiani, A., 1993. Tolerance and sensitization to the behavioural effects of drugs.
618 Behavioural Pharmacology 4, 289-312.
619
620 Stürzenbaum, S.R., Georgiev, O., Morgan, A.J., Kille, P., 2004. Cadmium detoxification in
621 earthworms: From genes to cells. Environmental Science and Technology 38, 6283-6289.
622
623 Suzuki, M.M., Bird, A., 2008. DNA methylation landscapes: provocative insights from
624 epigenomics. Nature Reviews Genetics 9, 465-476.
625
626 Takiguchi, M., Achanzar, W.E., Qu, W., Li, G.Y., Waalkes, M.P., 2003. Effects of cadmium on
627 DNA-(Cytosine-5) methyltransferase activity and DNA methylation status during cadmium-induced
628 cellular transformation. Experimental Cell Research 286, 355-365.
629
630 Thomas, D.J., Styblo, M., Lin, S., 2001. The cellular metabolism and systemic toxicity of arsenic.
631 Toxicology and Applied Pharmacology 176, 127-144.
632
633 van Straalen, N.M., Janssens, T.K.S., Roelofs, D.A.F.N.v.S.N.M., Janssens, T.K.S., Roelofs, D.,
634 2011. Micro-evolution of toxicant tolerance: from single genes to the genome's tangled bank.
635 Ecotoxicology 20, 574-579.
636
637 Van Straalen, N.M., Roelofs, D., 2007. An Introduction to Ecological Genomics. Oxford University
638 Press, Oxford, UK, 307 pp.

639
640 Vatamaniuk, O.K., Bucher, E.A., Sundaram, M.V., Rea, P.A., 2005. CeHMT-1, a putative
641 phytochelatin transporter, is required for cadmium tolerance in *Caenorhabditis elegans*. Journal of
642 Biological Chemistry 280, 23684-23690.
643
644 Vonlanthen, P., Bittner, D., Hudson, A.G., Young, K.A., Muller, R., Lundsgaard-Hansen, B., Roy,
645 D., Di Piazza, S., Largiader, C.R., Seehausen, O., 2012. Eutrophication causes speciation reversal in
646 whitefish adaptive radiations. Nature 482, 357-362.
647
648 Xiong, L.Z., Xu, C.G., Maroof, M.A.S., Zhang, Q.F., 1999. Patterns of cytosine methylation in an
649 elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification
650 polymorphism technique. Molecular and General Genetics 261, 439-446.
651
652 Zhao, C.Q., Young, M.R., Diwan, B.A., Coogan, T.P., Waalkes, M.P., 1997. Association of arsenic-
653 induced malignant transformation with DNA hypomethylation and aberrant gene expression.
654 Proceedings of the National Academy of Sciences of the United States of America 94, 10907-
655 10912.
656

657 Zhong, C.X.Y., Mass, M.J., 2001. Both hypomethylation and hypermethylation of DNA associated
658 with arsenite exposure in cultures of human cells identified by methylation-sensitive arbitrarily-
659 primed PCR. Toxicology Letters 122, 223-234.

660 **Table 1.** Summary data for trace element concentrations, soil pH and wt% loss on ignition (%LOI) for soil samples collected from sites across the
661 Devon Great Consols mine complex located in south-west England. For site locations see Fig. 1. Values are means of triplicate subsamples, standard
662 deviations are given in brackets.
663

	Al	**As**	Ba	Cd	Co	Cr	**Cu**	Fe	Mn	Ni	Pb	Zn	pH	%LOI
	mg/kg	**mg/kg**	mg/kg	mg/kg	mg/kg	mg/kg	**mg/kg**	mg/kg	mg/kg	mg/kg	mg/kg	mg/kg		
Site 1	9430	**4620**	47.2	0.7	19.5	10.8	**529**	48000	430	16.4	61	134	4.5	4.2
	(6720)	**(2020)**	(32.4)	(0.4)	(11.5)	(9.8)	**(266)**	(24600)	(335)	(15.8)	(35)	(74)		(0.1)
Site 2	6060	**5220**	70.0	< 0.2	52.7	6.9	**606**	99200	802	28.4	191	164	4.1	49.6
	(720)	**(470)**	(2.5)		(4.8)	(0.3)	**(27)**	(9300)	(96)	(2.8)	(21)	(22)		(0.4)
Site 3	17300	**6270**	45.9	0.18	25.7	17.8	**2647**	79600	630	22.3	225	277	4.8	29.7
	(3900)	**(1010)**	(8.4)	(0.09)	(5.4)	(3.6)	**(606)**	(2600)	(135)	(4.0)	(53)	(43)		(0.3)
Site 4	13600	**19200**	71.2	10.2	< 3.6	20.0	**910**	65900	262	9.5	148	63	4.6	17
	(2060)	**(3470)**	(8.9)	(1)		(3.3)	**(120)**	(10600)	(33)	(1.8)	(15)	(8)		(0.0)
SC	21500	**310**	45.5	0.41	< 3.6	31.7	**107**	45800	585	27.5	68	140	5.6	17.2
	1800	**(70)**	(2.8)	(0.4)		(4.2)	**(16)**	(6350)	(103)	(9.1)	(8)	(34)		(0.2)
OSC	7840	**<50**	37.0	< 0.2	< 3.6	10	**14**	1420	427	3.7	21	69	4.4	14.5
	(4770)		(20.7)			(6.0)	**(8)**	(8650)	(234)	(2.8)	(12)	(35)		(0.2)

664

FIGURE LEGENDS

Figure 1: Aerial images showing the location of the Devon Great Consols mine site in the South West UK (top right insert panel) and locations of the 5 sampling locations (Site 1-4 and Site SC) situated in the area on, and immediately adjacent to, the Devon Great Consols mine workings.

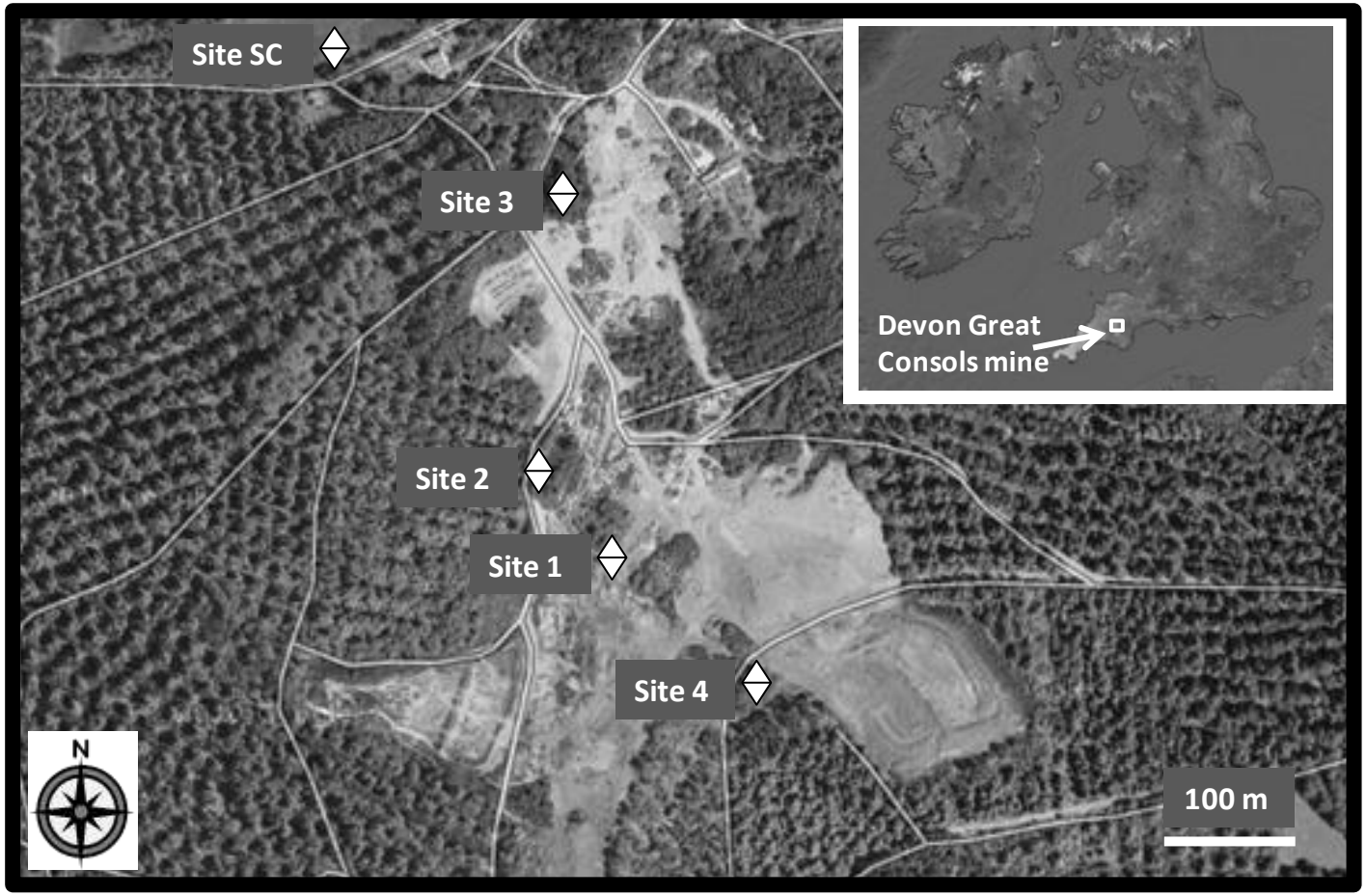
Figure 2: Temporal patterns of survival of *L. rubellus* collected at five locations of contrasting geochemistry (4 polluted and 1 site reference) within the Devon Great Consols mine complex and surrounding area and a known naïve population following exposure to 300 mg/kg of arsenic in a spiked clay loam soil over 14 days.

Figure 3: Mitochondrial and nuclear analysis of *L. rubellus* population structure and corresponding mitochondrial mismatch distributions of collected *L. rubellus*. Panel A: shows a phylogenetic tree of mitochondrial COII genotype showing branching of major lineage (Left and right hand branches of the network are denoted Lineage B & A respectively) and the numbers of individuals from each site within the lineages. Panel B: AFLP multi-locus profiling PCO analysis showing individuals from the six sample stations. Lineage A individuals cluster to the right on PC1, Lineage B to the left. Hybrids (found at SC and OSC only) lie between and above the two Lineage Groups.

Figure 4: Nuclear genome analysis of *L. rubellus* collected at six sites (4 polluted and 2 reference) of contrasting geochemistry within the Devon Great Consols mine complex and surrounding area. Panel (i) shows the result of a PCO of AFLP profiles for *L. rubellus* unambiguously ascribed to Lineage A, Panel (ii) shows the result of a PCO of AFLP profiles for *L. rubellus* unambiguously ascribed to Lineage B, Panel (iii) shows the result of a PCO of methylation sensitive AFLP analysis of *L. rubellus* unambiguously ascribed to Lineage A, Panel (iv) shows the result of a PCO of methylation sensitive AFLP analysis of *L. rubellus* unambiguously ascribed to Lineage B.

690 **Fig. 1.**

691



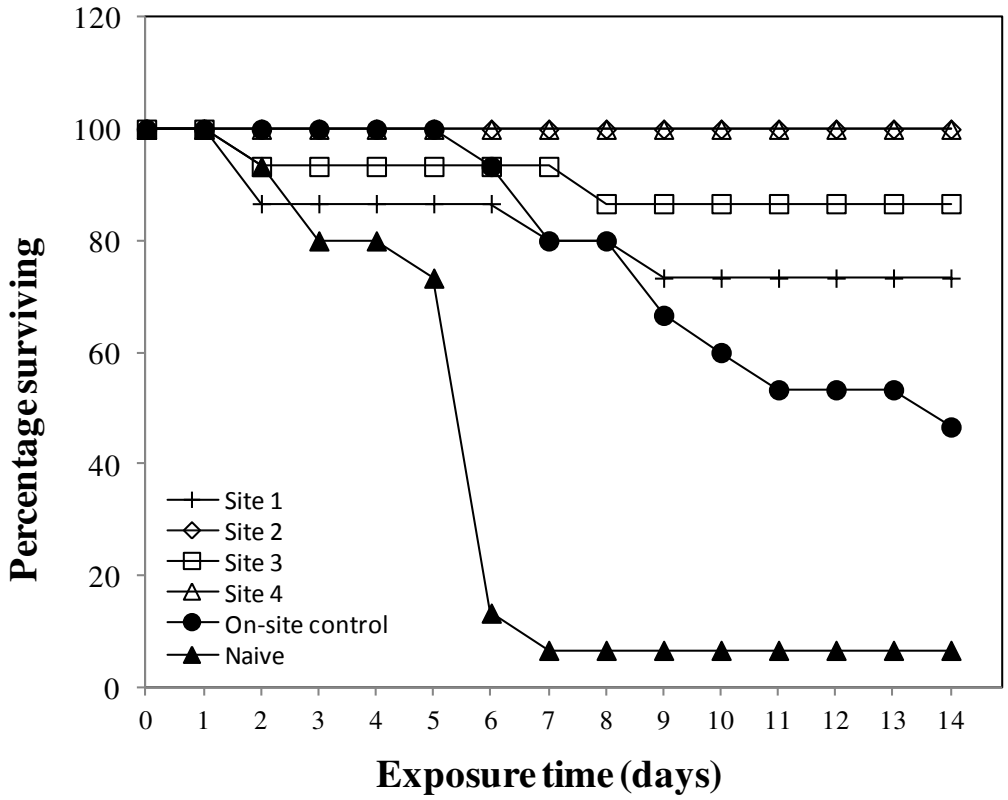
692

693

694 **Fig. 2.**

695

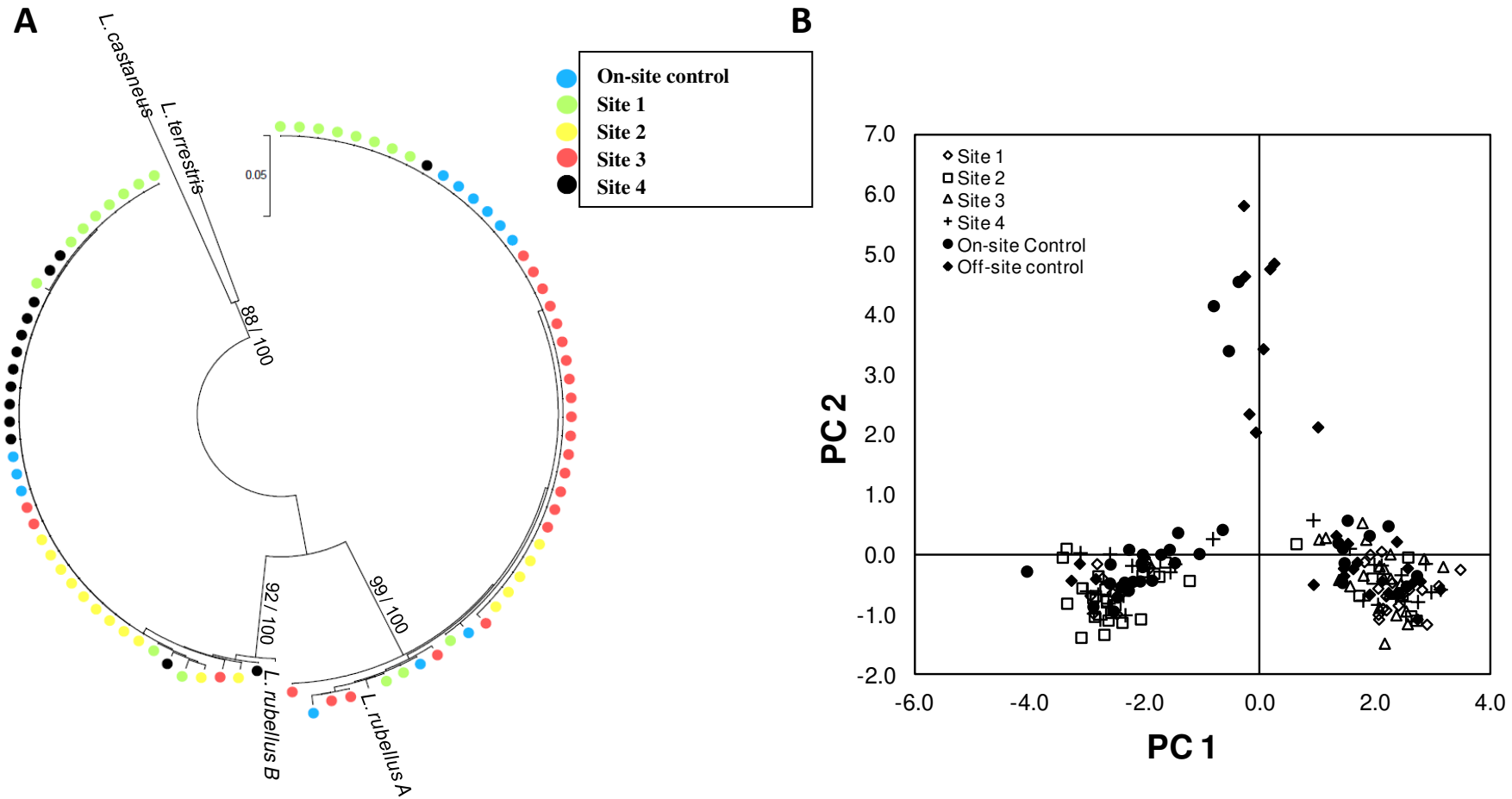
696



697

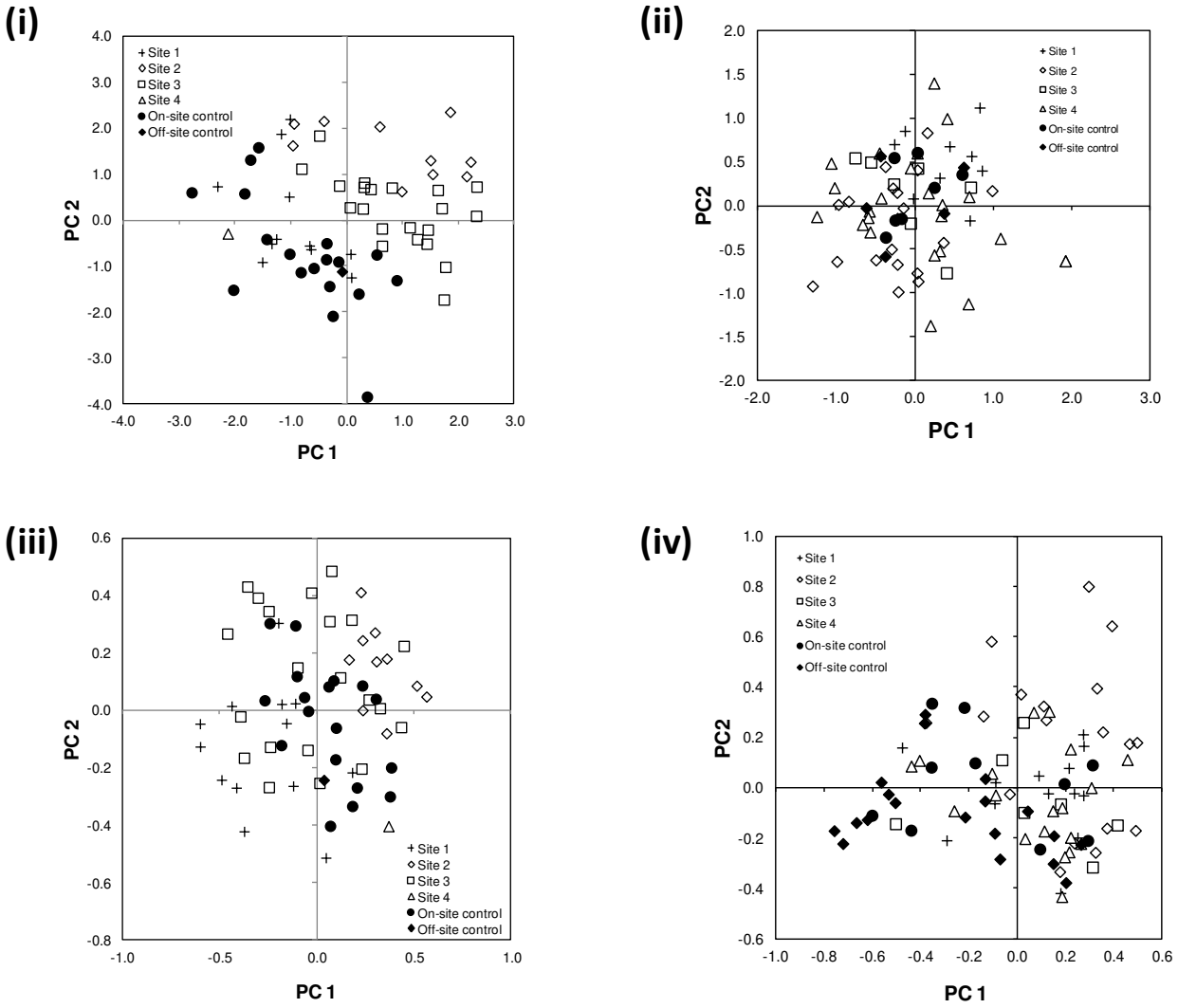
698 **Fig. 3.**

699



700

701 **Fig. 4.**


(i)
Site 1
Site 2
Site 3
Site 4
On-site control
Off-site control
PC 2
PC 1
(ii)
PC2
PC 1
(iii)
PC 2
PC 1
(iv)
PC2
PC 1

702